DOCUMENT RESUME

ED 187 225 HE 012 691

AUTHOR Braskamp, Larry A.; And Others
TITLE Student Written Comments: Dimensions of Instructional Quality.
PUB DATE Apr 80
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Attitude Measures; College Faculty; College Students; *Course Evaluation; Course Organization; *Faculty Evaluation; Higher Education; *Questioning Techniques; *Questionnaires; *Student Evaluation of Teacher Performance; Teacher Attitudes; Teacher Characteristics; Teaching Skills
IDENTIFIERS *University of Illinois Urbana Champaign

ABSTRACT

In a study at the University of Illinois, student written comments to four open-ended questions in end-of-semester student evaluations of 60 courses from ten different fields, taught by instructors of different ranks, were analyzed and classified into 22 categories. Nearly two-thirds were about the instructor, and one of four comments were on teaching skills, especially the ability to communicate effectively. The next most frequent comment was on knowledge of the field, followed by humanistic and personal characteristics of the instructor. The characteristics that students choose for evaluation differ from those faculty choose; the latter focuses more on course management than on teaching skills or personal qualities. Written comments and objective item responses were found to correlate, although more highly for comments about the instructor than for those about the course. It is suggested that open-ended evaluation questions be used regularly on these questionnaires. A list of references, a frequency distribution chart, and response groupings and correlations are appended. (MSE)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED187225

# STUDENT WRITTEN COMMENTS:

# DIMENSIONS OF INSTRUCTIONAL QUALITY

Larry A. Braskamp
John C. Ory
David M. Pieper

University of Illinois

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John Ory

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Requests for reprints should be sent to: Larry A. Braskamp
Measurement and Research Division
307 Engineering Hall
University of Illinois
Urbana, IL 61801

HE 012 691

## Abstract

Student written comments to four open-ended questions used in end-of-semester student evaluations of 60 courses from ten different fields, taught by instructors of different ranks, were content analyzed and classified into 22 dimensions. Nearly two thirds were about the instructor, and one of four instructor comments pertained to pedogogical skills. The favorableness of written comments about the instructor and course were sufficiently convergent with ratings based on fixed alternative items measuring overall instructor performance and course quality. While similar impressions of overall quality are provided by student responses to open-ended and fixed-alternative items, it was suggested that each type of evaluative information provides useful and different information.

## Introduction

One major method for the empirical study of teaching, especially instruction at the college level, has been the investigation of teacher traits and characteristics (Peterson and Walberg, 1979). Teaching effectiveness is viewed as a consequence of the personality and skills of the teacher. This approach has been extensively used in the development of student surveys and is now used for evaluating instructor competence. Many of the survey items have been derived from comments about effective teachers solicited from students, faculty and alumni (e.g., Hildebrand, 1973; Wilson, Dienst & Watson, 1973; Crawford and Bradshaw, 1968).

In the past decade numerous factor analytic studies have been run on sets of items assumed to have included the major dimensions of effective instruction identified in the early studies. The factors, which vary from study to study (e.g., Doyle & Whitely, 1974; Linn, Centra & Tucher, 1975) usually include clusters of items with the labels: instructor pedogogical skill, teacher student rapport, course organization and structure, course difficulty, and student outcomes. More recent work has focused on the dimensions of the ideal or best teacher. Feldman (1976) drew a composite view of the ideal teacher by synthesizing 23 research studies in which factors were correlated with overall competence. Stimulation was the most highly correlated dimension with overall teaching ability; enthusiasm, knowledge of subject matter, preparation, organization of material, and clarity were also highly associated with overall teaching ability. Based on student and faculty comments for nominations of outstanding teachers, Goldsmid, Gruber, and Wilson (1977) identified two general traits: competence (command of the material) and conscientiousness (concern for students' achievement.)

However, there has not been an analysis of written comments of students who complete a rating form administered during the latter part of a course. Do students write comments that encompass the characteristics identified in factor analytic studies and those associated with excellence in teaching? The first purpose of this study was to investigate the types of comments students make in an end-of-course evaluation and to compare them with those identified in the earlier studies.

The second major purpose of the study was to determine the congruent validity of student written comments and overall ratings of instructor and course quality as measured by fixed alternative items. Several studies on the convergent and discriminant validity of student and instructor self ratings have been conducted

(Doyle & Crichton, 1978; Braskamp, Caulley, & Costin, 1979) but only one study involving different methods of data collection from the same source (student) has been identified. Judgments of overall quality of the instructor and the course based on student ratings, written comments to open-ended questions and student interviews were highly similar (Ory, Braskamp, & Pieper, in press).

By using different multiple source and/or multiple method approaches, triangulation of the measurement process (Webb, Campbell, Schwartz & Sechrest, 1971) and structural corroboration are possible (Eisner, Note 3). The second purpose of this study was to further investigate the corroboration between the favorableness of student written comments and of student ratings to fixed alternative (objective) items.

Specifically, two questions were asked in this study:

1. What dimensions underlie student comments?
2. Do written comments and fixed alternative ratings reveal similar degrees of favorableness/unfavorableness?

## Method

### Sample

Of the 2,685 courses evaluated during the fall 1978 semester in which the Instructor and Course Evaluation System (ICES) was used (Office of Instructional Resources, Note 2), 60 courses were randomly selected given the following three restrictions. Courses were initially classified by HEGIS code into ten groups according to departmental affiliation: Agriculture, Architecture, Business, Chemistry/Biology, Classics, Education, Fine Arts, Math, Physical Sciences, and Social Sciences. Second, courses with an estimated enrollment of between 10 and 50 students were selected for the study. Third, six instructors selected from each group included two professors and one teaching assistant or instructor teaching upper undergraduate or graduate courses and one professor and two teaching assistants teaching lower undergraduate courses.

### Instrument

The ICES student rating questionnaire, completed by students during the final weeks of each semester, is used as part of the cafeteria student rating system that has been in operation for two years. On all questionnaires, students complete six demographic items and three evaluation items: "Rate the course," "Rate the instructor," and "Rate the course content." Students indicate their response to the three items on a six point scale anchored by "poor" and "excellent." Instructors are given the opportunity to select the remaining 23 items that they consider most appropriate for their course and informational needs. On side two of the ICES form, students are asked to write their comments to four questions: "What are the major strengths and weaknesses of the instructor?" "What aspects of this course were most beneficial to you?" "What do you suggest to improve this course?" and "Comment on the grading procedures and exams." Spaces for two instructor selected questions are provided.

## Data Classification and Analysis

For courses in which more than 25 student questionnaires were completed, a random sample of 25 questionnaires was selected. For each course, the following summary data were recorded: the level of the course (100 through 400); the rank of the instructor; the number of students completing questionnaires; and the percent of the students who expected a grade of A, indicated the course was required, and were juniors, seniors or graduate students. Average ratings on the three global fixed alternative items of these students were also recorded. Three judges independently read samples of written comments to develop content categories of student opinion. The classification scheme used in the study is a result of merging the three independent classification schemes.

## Instructor

1. Attitude--interest in, enthusiasm for, and attitude toward teaching, course, and subject matter.

2. Rapport--concern and care for individual students; warmth, understanding; knows student names; bias toward a few students.

3. Person--professional, nervous, impatient, amiable personality; has annoying mannerisms; sincere; expresses humor, self confident.

4. Knowledgeable--knows subject matter, material; well versed.

5. Stimulation--challenging, fascinating; creates interest in field, stimulates thought; encourages new ideas; boring.

6. Rigor--expects too much, too little; challenges class to improve; poor discipline in class.

7. Openness--flexible; open to differing and opposite ideas; self righteous.

8. Ability--makes clear and concise presentations; talks too fast, too slow; easy to take notes; explains well; uses good examples.

9. Preparedness--follows an outline; poorly prepared for class; good organization of lecture.

10. Helpfulness--patient; helpful; readily available; spends time with individual students outside class.

11. Teacher overall--excellent teacher, good, poor role model; desire to take more courses.

12. Improvement--instructor became more relaxed, better lecturer at end of semester.

13. Miscellaneous--noncodable responses.

Course

14. Organization--clarity of expectations, objectives; pace was even, uneven; good, poor course structure; lectures and labs were integrated.

15. Content--learned subject matter; learned writing, speaking, calculating, selling skills.

16. Materials--textbooks, handouts, audio visual materials were good, poor, used effectively; lab equipment.

17. Requirements--too many, too few required assignments; appropriate projects; vague assignments.

18. Lecture--lectures are good, poor, not worth attending, integrative.

19. Discussion--class interactions are good, relevant; discussions are interesting, superficial; little accomplished in discussions.

20. Specific aspects--classroom laboratory, studio facilities; time of day of course.

21. Overall course--enjoyment of course; good course; course is beneficial; course opens up new ideas, stimulates my thinking.

Grading

22. Fairness; too easy/too strict; ambiguous test questions; changed grading procedures during semester; unclear grading procedures.

A total of 3240 different student comments from 924 students in the 60 courses were classified into one of the 22 dimensions. All comments were classified by three judges, who were staff members in a campus office of instructional resources. When the three judges disagreed on the placement of a comment, they discussed the comment until a consensus of the placement was obtained. Nineteen or 1.2 percent of the comments about the instructor were non-codable.

Each student's comments were read and classified into as many of the 22 dimensions as needed. A value of three was assigned to a dimension if all the students' comments classified into that dimension were positive, a value of two if both positive and negative comments were made, and a value of one if only negative comments were written. If a student made no comments classified into a dimension, no value was assigned to the dimension.

## Results

Data analyses and results are organized around the two questions asked in this study.

What dimensions underlie student comments? The 3240 comments classified by degree of favorableness and dimension are presented in Table 1. Nearly two thirds of all comments were positive (62.6%). One of every two comments was about a

characteristic of an instructor, whereas only one of every three was about a dimension of the course even though two of the four questions were related to the course. Of the 1607 comments about the instructor, nearly one fourth were about the instructor's pedagogical abilities (see dimension 1 in Table 1). The second most common characteristic was instructor knowledge of field followed by three characteristics reflecting instructor personality and relationship with students--instructor as a person, rapport, and helpfulness.

Over one third of the student comments about the course related to student learning of the course content. Student comments about materials, specific aspects of the course, assignments and workload accounted for an additional forty percent of all comments. The organization of the course (pace, integration, clarity of objectives) was commented on less than nine percent of the time.

Do written comments and fixed alternative ratings reveal similar degrees of favorableness/unfavorableness? Table 2 presents the number and percent of comments classified as either positive or negative, by content area, and by the objective rating, "Rate the instructor." Within each content area, the relative frequencies of positive and negative comments do differ according to objective rating (all three $\chi^2$ statistics were significant, $\alpha = .01$). Students who rated the instructor a 6 on the six-point scale had disproportionately more positive comments than negative whereas those students who gave ratings of 3 or below had more negative than positive comments. Student evaluations of the instructor most clearly show the congruency between the two data collection methods, but this can be expected since the objective rating was on the instructor. The congruency between the two data collection methods was further examined through statistical analysis in which the course was the unit of analysis. Average ratings on the two objective items and an overall favorableness rating of the course and instructor based on the written comments were calculated for each course. The overall favorableness ratings of the instructor were calculated in which the weights assigned (3 = positive, 2 = positive/negative, and 1 = negative) to all comments made about the instructor were summed and then divided by the number of comments made. The same calculations were made about the course. Thus there were two global ratings for both the instructor and the course, one based on the objective fixed alternative items and the other on the written comments. Table 3 shows the correlations among the four global indicators of quality. The convergence between the two ratings of the same trait (teacher, course) by two different data collection methods (items, comments) is more apparent for the instructor ratings ($r = .75$) than for the course ratings ($r = .50$). Student ratings of the instructor and course based on the objective items was considerably more highly interrelated ($r = .75$) than the ratings based on the written comments ($r = .26$).

## Discussion

Student comments from open-ended questions about the instructor, course aspects, suggestions for improvement and grading show that students offer a variety of comments. However, students concentrate their comments on only a few dimensions, with two of three comments being about an instructor characteristic. Based on the content analysis, students focus more frequently on the instructor's pedagogical skills, the ability to clearly communicate. The next most frequently mentioned dimension was instructor knowledge of the field followed by more humanistic and personal qualities of the instructor.

The diversity of written comments equals the broad range of instructional dimensions identified in most factor analytic studies. As a consequence, the classification scheme developed in this study of student ratings includes most, if not all, identified factors. On the other hand, the profile of characteristics that students used to evaluate the teacher and the course differ slightly from earlier studies describing ideal teachers or nominations of outstanding teachers. Earlier studies emphasized two dimensions of instructional quality: the interpersonal and pedogogical skills of the instructor. In our analysis, relatively more comments were directed toward the latter of the two skill areas. More comments were directed toward instructor ability to communicate and instructor knowledge of the field and relatively less on the conscientiousness or enthusiasm of the instructor. Also, the instructor's ability to stimulate student learning was more emphasized in student descriptions or nominations of excellent teachers than in the post-course ratings of this study. When students are routinely evaluating their instructors, they perhaps do not focus as much on the concern the instructor had for them as individuals as when they are asked to describe good teaching and good teachers.

The characteristics that students used to evaluate instruction at the end of the course are also different from those that faculty select to evaluate their instruction. A study of faculty selection of evaluation items included on post-course rating forms revealed that instructors focused more on the management of their course than on their teaching skills and personal qualities (Ory, Brandenburg & Pieper, in press). Faculty are therefore encouraged to include open-ended questions along with selected fixed alternative items to collect student opinion on topics otherwise overlooked.

A high degree of convergent validity between written comments and objective item responses was obtained. However, students' objective ratings and written comments about the instructor were more highly correlated than were students' objective ratings and written comments about the course in general. The reason for this discrepancy is not known. Students may be able to rate more uniformly on instructor than on course or a halo effect that some researchers argue which pervades overall impressions of an instructor (Feldman, 1976) is strong enough to influence both types of ratings. This congruency does, however, have practical implications. An average class rating on a fixed alternative item of overall quality of the instructor is more cost effective than a content analysis of student written comments if overall instructor competence, as judged by students, is a desired piece of evaluative information. The similarity between overall impressions should not, however, discourage the collection and analysis of student written comments. Student responses to open-ended questions provide different and more specific evaluative information than do average class ratings. Furthermore, instructors often find the former information more useful and trustworthy for improving their own teaching (Office of Instructional Resources, Note 3).

Students' written comments and responses to objective items provide highly similar ratings of overall instructional quality. It would appear that over the years the development of student rating forms has come full circle. Many of today's computer analyzed rating scale items were developed from the content identified through written student comments about instructors and courses. Therefore, it should not be surprising that students respond in similar ways to one type of item (fixed alternative) that is derived from another type of item (the open-ended question).

## Reference Notes

1. Eisner, E. W. The perceptive eye: Toward the reformation of educational evaluation. Occasional papers of the Stanford Evaluation Consortium, Stanford University, Stanford, California. Mimeographed, December, 1975.

2. ICES: Its rationale and description, ICES Newsletter No. 2. Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign. Mimeographed, August, 1977.

3. Summary of Research on ICES, ICES Research Report No. 1. Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign. Mimeographed, February, 1980.

## References

Braskamp, L. A., Caulley, D. & Costin, F. Student ratings and instructor self-ratings and their relationship to student achievement. American Educational Research Journal, 1979, 16, 295-306.

Crawford, P. L. & Bradshaw, H. L. Perception of characteristics of effective university teachers: A scaling analysis. Educational and Psychological Measurement, 1968, 28, 1079-1085.

Doyle, K. O., Jr. & Whitely, S. E. Student ratings as criteria for effective college teaching. American Educational Research Journal, 1974, 11, 259-274.

Doyle, K. O. & Crichton, L. I. Student, peer and self evaluation of college instructors. Journal of Educational Psychology, 1978, 70, 815-826.

Feldman, K. A. The superior college teacher from the students' view. Research in Higher Education, 1976, 9, 199-242.

Goldsmid, C.A., Gruber, M.E. & Wilson, E.K. Perceived attributes of superior teachers (PAST): An inquiry into the giving of teacher awards. American Educational Research Journal, 1977, 14, 423-440.

Hildebrand, M. The character and skills of the effective professor. Journal of Higher Education, 1973, 44, 41-50.

Linn, R. L., Centra, J. A. & Tucher, L. R. Between, within, and total group factor analyses of student ratings of instruction. Multivariate Behavioral Research, 1975, 10, 227-288.

Ory, J. C., Brandenburg, D. C. & Pieper, D. M. Selection of course evaluation items by high and low rated faculty. Journal of Research in Higher Education, in press.

Ory, J. C., Braskamp, L. A. & Pieper, D. M. The congruency of student evaluative information collected by three methods. Journal of Educational Psychology, in press.

Peterson, P. L. & Wallberg, H. J. (eds.), Research on Teaching. Berkeley, CA: 1979.

Webb, E. J., Campbell, D. T., Schwartz, R. D. & Sechrest, L. Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago, Illinois: Rand McNally & Co., 1971.

Wilson, R. C., Dienst, E. R. & Watson, N. L. Characteristics of effective teachers as perceived by their colleagues. Journal of Educational Measurement, 1973, 10, 31-37.

Table 1

Frequency Distribution of Student Written Comments

Classified by 22 Evaluative Dimensions

| Dimension | Comments: Positive f | Positive % | Positive + Negative f | Positive + Negative %[a] | Negative f | Negative %[a] | Total f | Total %[b] |
|---|---|---|---|---|---|---|---|---|
| About Instructor | | | | | | | | |
| Ability | 176 | 19.1 | 38 | 4.1 | 176 | 19.1 | 390 | 24.? |
| Knowledge | 196 | 21.2 | 6 | .17 | 17 | 1.8 | 219 | 13.6 |
| Person | 141 | 15.3 | 7 | .8 | 39 | 4.2 | 187 | 11.6 |
| Rapport | 116 | 12.6 | 1 | .1 | 34 | 3.7 | 151 | 9.4 |
| Helpfullness | 125 | 13.5 | 2 | .2 | 10 | 1.1 | 137 | 8.5 |
| Preparedness | 69 | 7.5 | 3 | .3 | 54 | 5.9 | 126 | 7.8 |
| Teacher Overall | 106 | 11.5 | 1 | .1 | 5 | .5 | 112 | 7.0 |
| Stimulating | 65 | 7.0 | 0 | 0 | 27 | 2.9 | 92 | 5.7 |
| Attitude | 64 | 6.9 | 1 | .1 | 19 | 2.1 | 84 | 5.2 |
| Openness | 28 | 3.0 | 0 | 0 | 14 | 1.5 | 42 | 2.6 |
| Rigor | 10 | 1.1 | 1 | .1 | 28 | 3.0 | 39 | 2.4 |
| Miscellaneous | 0 | 0 | 0 | 0 | 19 | 2.1 | 19 | 1.2 |
| Improvement | 10 | 1.1 | 0 | 0 | 0 | 0 | 10 | .6 |
| Total | | | | | | | 1608 | 99.9 |
| About Course | | | | | | | | |
| Content | 329 | 35.6 | 17 | 1.8 | 33 | 3.6 | 379 | 35.8 |
| Material | 67 | 7.3 | 8 | .9 | 75 | 8.1 | 150 | 14.2 |
| Specific Aspects | 69 | 7.5 | 0 | 0 | 69 | 7.5 | 138 | 13.0 |
| Requirements | 58 | 6.3 | 4 | .4 | 71 | 7.7 | 133 | 12.6 |
| Course Overall | 73 | 7.9 | 0 | 0 | 21 | 2.3 | 94 | 8.9 |
| Organization | 11 | 1.2 | 0 | 0 | 75 | 8.1 | 86 | 8.1 |
| Discussion | 35 | 3.8 | 1 | .1 | 4 | .4 | 40 | 3.8 |
| Lecture | 12 | 1.3 | 3 | .3 | 23 | 2.5 | 38 | 3.6 |
| Total | | | | | | | 1058 | 100 |
| About Grading | | | | | | | | |
| Grading | 269 | 29.1 | 80 | 8.7 | 225 | 24.4 | 574 | 100 |

[a]Percent of the 923 students indicating specified content and degree of favorableness

[b]For dimensions 1-13, the percent of all written comments about the instructor. For dimensions 14-21, the percent of all written comments about the course. For dimension 22, the percent of all written comments about grading.

Table 2

Frequency and Percent of Positive and Negative Student Written Comments Grouped by Three Major Content Areas and Objective Ratings of Instructor

| | Objective Rating | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 6 | | 5 | | 4 | | 3, 2, & 1 | |
| | f | % | f | % | f | % | f | % |
| **About Instructor** | | | | | | | | |
| Positive comments | 455 | 88.5 | 388 | 75.9 | 148 | 60.2 | 105 | 39.3 |
| Negative comments | 59 | 11.5 | 123 | 24.1 | 98 | 39.8 | 162 | 60.7 |
| $\chi^2 = 60.17$* | | | | | | | | |
| **About Course** | | | | | | | | |
| Positive comments | 215 | 76.0 | 254 | 73.6 | 103 | 63.6 | 82 | 45.8 |
| Negative comments | 68 | 24.0 | 91 | 26.4 | 59 | 36.4 | 97 | 54.2 |
| $\chi^2 = 24.78$* | | | | | | | | |
| **About Grading** | | | | | | | | |
| Positive comments | 97 | 70.8 | 116 | 58.0 | 32 | 42.7 | 24 | 29.3 |
| Negative comments | 40 | 29.2 | 84 | 42.0 | 43 | 57.3 | 58 | 70.7 |
| $\chi^2 = 39.12$* | | | | | | | | |

*Crit$\chi^2_{.01} = 11.34$ Chi square statistic computed on frequency percentages.

Table 3

Correlations Between Mean Instructor and Course Ratings Based on Student Written Comments and Objective Item Responses (n=60)

| | Written Comments | | Objective Ratings | |
|---|---|---|---|---|
| | Instructor | Course | Instructor | Course |
| | 1 | 2 | 3 | 4 |
| 1 | ____ | | | |
| 2 | .26 | ____ | | |
| 3 | .75* | .44* | ____ | |
| 4 | .43* | .50* | .75* | ____ |

*p < .01